REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
26-05-2015	Final Report	25-Aug-2009 - 24-Oct-2014

4. TITLE AND SUBTITLE
Final Report: Strategic State Estimation in Uncertain and Mixed Multiagent Environments

5a. CONTRACT NUMBER
W911NF-09-1-0464

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
Prashant Doshi (PI), Adam Goodie (Co-PI), Dan Hall (Co-PI), Roi Ceren and Matthew Miesel

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Georgia Research Foundation,
200 D. W. Brooks Drive

Athens, GA 30602 -5016

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
55749-NS.7

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
Human judgment of uncertainty suffers from various cognitive biases. In the military context, operators of autonomous vehicles such as unmanned aerial vehicles (UAVs) may assess probabilities at levels that are not objectively justified. Some of the psychological issues that affect probability judgment include inaccuracy in verbally reporting probabilities and a lack of honesty in expressing the true probability assessments. The objectives of this project were to take bold and critical steps toward establishing field-valid probability assessment and update techniques using a realistically simulated UAV setting. Furthermore, the project studied and computationally

15. SUBJECT TERMS
Probability judgment; ambiguity; Brier score; cognition

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Prashant Doshi
UU	UU	UU	UU		19b. TELEPHONE NUMBER: 706-583-0827

Report Title

Final Report: Strategic State Estimation in Uncertain and Mixed Multiagent Environments

ABSTRACT

Human judgment of uncertainty suffers from various cognitive biases. In the military context, operators of autonomous vehicles such as unmanned aerial vehicles (UAVs) may assess probabilities at levels that are not objectively justified. Some of the psychological issues that affect probability judgment include inaccuracy in verbally reporting probabilities and a lack of honesty in expressing the true probability assessments. The objectives of this project were to take bold and critical steps toward establishing field-valid probability assessment and update techniques using a realistically simulated UAV setting. Furthermore, the project studied and computationally modeled validated static probability assessments and probability updates. The research was performed by a multi-disciplinary team consisting of faculties in computer science, cognitive psychology and statistics. Five psychological studies investigated the presence of cognitive biases in judgment and techniques to obtain valid probability assessments in a laboratory setting.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Adam Goodie, Prashant Doshi, Daniel Hall, Mathew Meisel, Roi Ceren and Erica Fortune, Strategic State Estimation in Uncertain and Mixed Multiagent Environments. Poster, Annual Meeting of the Society for Judgment and Decision Making (SJDM), Minneapolis, MN, Nov, 2012.

Adam Goodie, Prashant Doshi, Daniel Hall, Roi Ceren and Matthew Miesel, The Role of Incentive Schemes in Probability Assessment Under Uncertainty, US Army Conference on Applied Statistics (ACAS), Monterrey, CA, Oct 26, 2012.

Prashant Doshi, Adam Goodie, Daniel Hall, Matthew Meisle, and Roi Ceren, Evaluating the Validity of Probability Assessments in Strategic and Realistic Environments, US Army Conference on Applied Statistics (ACAS), Annapolis, MD, Oct 20, 2011.

Number of Presentations: 3.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper	
08/07/2013	4.00	Roi Ceren, Prashant Doshi, Matthew Meisel, Adam Goodie, Dan Hall. On Modeling Human Learning in Sequential Games with Delayed Reinforcements, IEEE International Conference on Systems, Man and Cybernetics (SMC). 13-OCT-13, . : ,

TOTAL: **1**

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received	Paper	
05/23/2015	6.00	Adam Goodie, Prashant Doshi, Mathew Miesel, Roi Ceren, Dan Hall. Evaluating and Improving Probability Assessment in an Ambiguous, Sequential Environment, Current Psychology (10 2014)
08/16/2013	5.00	Adam Goodie, Mathew Meisel, Roi Ceren, Dan Hall, Prashant Doshi. Improving Probability Assessment in a Sequential and Ambiguous Environment, Judgement and Decision Making (09 2013)

TOTAL: **2**

Number of Manuscripts:

Books

Received Book

TOTAL:

Received Book Chapter

TOTAL:

Patents Submitted

Patents Awarded

Awards

Creative Research Medal, University of Georgia, 2011 to Prashant Doshi
Creative Research Medal, University of Georgia, 2012 to Adam Goodie

Graduate Students

NAME	PERCENT_SUPPORTED	Discipline
Roi Ceren	0.33	
Mathew Miesel	0.33	
Erica Fortune	0.04	
Xia Qu	0.12	
FTE Equivalent:	**0.82**	
Total Number:	**4**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Prashant Doshi (PI)	0.08	
Adam Goodie (Co-PI)	0.13	
Daniel Hall (Co-PI)	0.04	
FTE Equivalent:	**0.25**	
Total Number:	**3**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED	Discipline
Amanda Falls	0.01	Computer and Computational Sciences
Ryan Gell	0.01	Computer and Computational Sciences
FTE Equivalent:	**0.02**	
Total Number:	**2**	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period:	2.00
The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:	2.00
The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:	0.00
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):	2.00
Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:	0.00
The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense	0.00
The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:	0.00

Names of Personnel receiving masters degrees

NAME	
Mathew Miesel	
Total Number:	**1**

Names of personnel receiving PHDs

NAME	
Xia Qu	
Total Number:	**1**

Names of other research staff

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

Approach

We investigated the presence of cognitive biases in judgment and techniques to obtain valid probability assessments using five psychological studies in a laboratory setting. Studies in this project involve participants from the general pool and those with specialized military knowledge. While the former type of participant is available from UGA's Psychology Research Pool, the latter is available from the Army and Air Force ROTC situated within UGA.

Studies 1-3 took place in the context of the Georgia test bed for autonomous control of vehicles (GaTAC). GaTAC is a computer simulation framework for evaluating autonomous control of aerial robotic vehicles such as UAVs (see Fig. 1 in attached Powerpoint file). It provides a low-cost and open-source laboratory alternative to highly complex and expensive simulation infrastructures. GaTAC deploys multiple instances of the open-source flight simulator, FlightGear, utilizing hyper-realistic 3D terrain data on a networked cluster of computing platforms. GaTAC provides a flight dynamics software module that translates high-level navigational actions to commands for the flight control surfaces. GaTAC will allow for maximum ecological validity of the studies in a laboratory setting. Previous investigations into biases in judgment utilized simple settings in which the relevant probabilities were few, were given and the calculations were often simple. GaTAC is developed in the THINC lab in the department of computer science at UGA.

Study 1 investigated whether participants' verbal probability assessments are unreliable and sought ways to validate the assessments. The primary alternative to direct verbal expressions is to determine preferences between betting on the outcome of the predictive probability event and on events with clear, objectively specified probabilities. The study will investigate whether the verbal reports are consistent with degrees of uncertainty as inferred from the choice data.

Study 2 investigated possible lack of honesty of general pool participants in reporting true probability assessments and sought ways to correct it. It examined whether proper scoring rules are needed to obtain honest and reliable reports of subjective probabilities.

Study 3 investigated possible lack of honesty of ROTC pool participants in reporting true probability assessments and sought ways to correct it. It examined whether proper scoring rules are needed to obtain honest and reliable reports of subjective probabilities.

Study 4 examined the effect of proper scoring rules on the general pool participants in the context of training interventions. The protocol did not use GaTAC and involved showing pre-defined trajectories to the participants as Powerpoint slides.

The final Study 5 analyzed the effect of two independent variables, use of a proper scoring rule such as the Brier scoring rule and training intervention, on probability judgments in our strategic game setting. The protocol did not use GaTAC and involved showing pre-defined trajectories to the participants as Powerpoint slides. Participant briefing was improved to include a demonstration of how the Brier scoring rule operated for improved understandability.

Results

We have completed running and analyzing all five studies during the grant period. We recruited a total of 467 participants for the studies, out of which 28 were from the ROTC pool while the remaining were from the general Psychology research pool. In Study 1 we did not observe any systematic inflationary or deflationary bias in the uncertainty expressions of the participants. However, the significantly smaller interval for ROTC participants despite the small sample indicates a much better behaved population in the context of this experiment, as compared to the general research pool. In the context of Study 2, which utilized predefined trajectories of the subject's UAV, we observed that monetary incentivization using a non-proper scoring rule, 0-1 scoring, resulted in a significant increase in payout and consequently better calibrated probability assessments. However, use of a proper scoring rule, Brier scoring, did not show similar benefit. We did not observe a main effect of any of the scoring rules for the ROTC pool in Study 3. Consequently, we did not replicate the positive finding of a main effect of 0-1 scoring on probability assessments as observed in Study 2. We did not observe a significant effect of the proper scoring rule among the participants in Study 4 despite the use of interventions. The results of Study 5 demonstrated the capacity of probability assessments to fall under the control of incentives, which are well understood, even under conditions of extreme ambiguity. We observed a statistically significant main effect of the proper Brier scoring (see Fig 2. in attached Powerpoint and table below). However, the main effect of intervention and the interaction effect between intervention and Brier scoring are not significant.

Main ANOVA results of the study.

	df	Type III SS	Mean Square	F value	p	partial • eta^2
Intervention	1	0.01062	0.01062	0.61	0.44	.006
Brier_Score	1	0.13573	0.13573	7.83	0.006**	.065
Int x Brier	1	0.00925	0.00925	0.53	0.47	.005
Residual	112	1.94153	0.01734			

Computational modeling of assessment data

An analysis of trend in participants judging the probabilities of successfully reaching the target in Study 4 is indicative of a learning effect across decision points and a learning effect as trials progress as well. Consequently, we are interested in computationally modeling this learning using a process-oriented, generative model with psychological plausibility.

Our results demonstrated that descriptive reinforcement learning with cognitive biases gets us close to modeling human judgments in contexts with delayed reinforcements but shows room for further improvement (see Fig. 3 in attached Powerpoint). Certain behaviors are challenging to computationally model, such as participants dropping their assessments in the later stages. This observation illuminates a pitfall of reinforcement learning: model assessments may propagate too slowly to precisely match the data set, an observation that has precedence. While participants may quickly change their assessments, temporal difference learning requires several iterations before a dramatic change is visible.

Technology Transfer

As part of technology transfer outreach, PI, Prashant Doshi, gave a presentation titled, Evaluating the Validity of Probability Assessments in Strategic and Realistic Environments, at the US Army Conference on Applied Statistics (ACAS) held in Annapolis, MD, on October 19, 2011. Co-PI Adam Goodie gave a presentation titled, The Role of Incentive Schemes in Probability Assessment Under Uncertainty, at the US Army Conference on Applied Statistics (ACAS) held in Monterrey, CA, on October 26, 2012.

The PI provided brief technical consult to John Nierwinski Jr. from the US Army Materiel Systems Analysis Activity (AMSAA) on the topic of combining information from multiple subject matter experts.

Strategic State Estimation in Uncertain and Mixed Multiagent Environments (55749)

PI: Prashant Doshi (CS)
Co-PI: Adam Goodie (Psych)
Co-PI: Dan Hall (Stat)
University of Georgia

- Human judgment of uncertainty suffers from cognitive biases
  - Operators of vehicles such as UAVs may generally assess probabilities of their predictions at levels not objectively justified
  - Uncertainty in realistic settings is difficult to judge and not objectively quantified
- Existing approaches inapplicable
  - No consensus on key issues related to probability judgment
  - Domain-general approaches for valid probability assessments may not apply to complex military context
    - Prior studies used simplistic settings (not strategic)

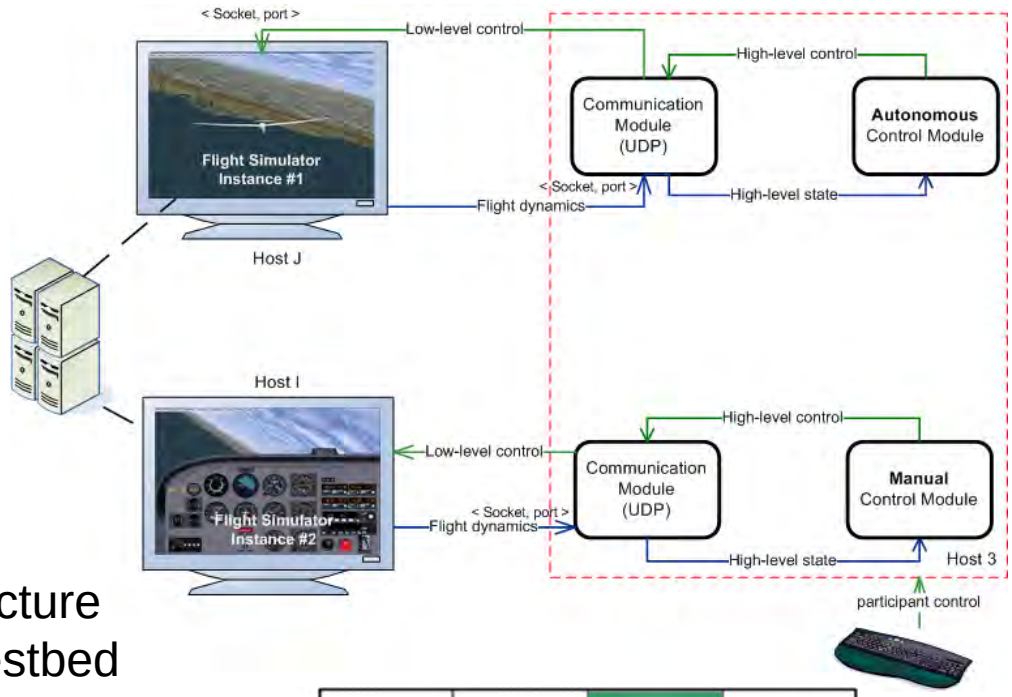

Figure 1: Architecture of the Georgia testbed for control of autonomous vehicles (GaTAC)

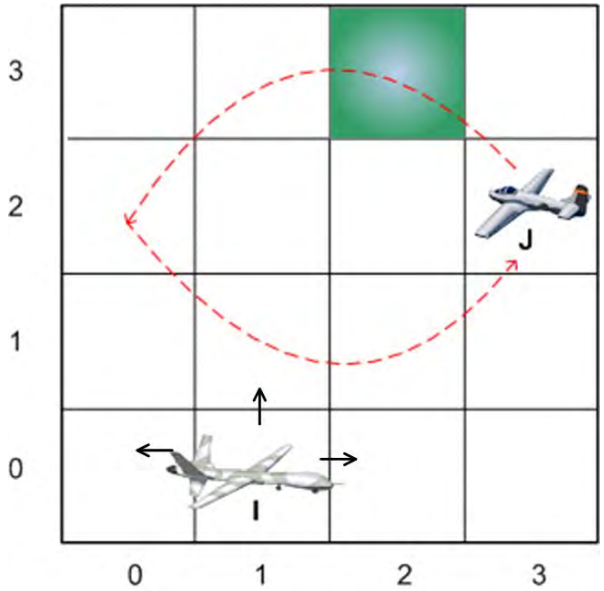

Figure 2. Average summed Brier score per trial in Study 5 with and without review interventions. We also compare the observed Brier scores with those that result from baseline assessments

	Brier-based incentive	No Brier-based incentive
Review	1.402	1.352
No Review	1.440	1.353
Random	1.233	
Point – optimal	1.274	
Point – nonresponse	1.262	

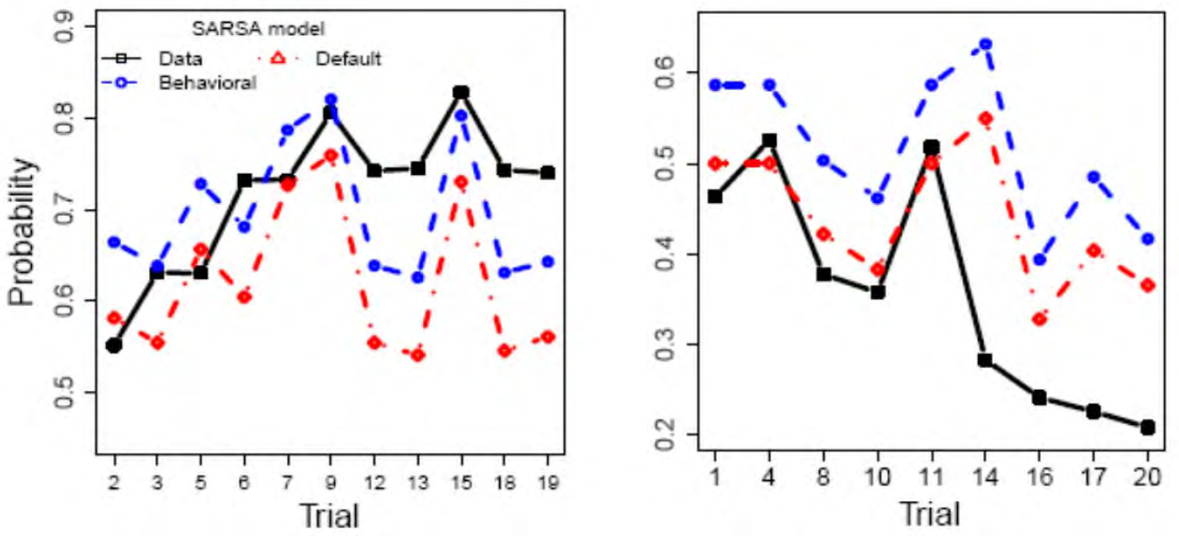

Figure 3. Average probability assessment in trials that led to wins (left) and losses (right) as observed in data and predicted by our computational models.